Postprint version of paper published in Food and Bioproducts Processing 2017, vol. 102, p. 1-19.
**The content is identical to the published paper, but without the final typesetting by the publisher.**

# Application of a dynamic metabolic flux algorithm during a temperature-induced lag phase

D. Vercammen[a], D. Telen[a], P. Nimmegeers[a], A. Janssens[a], S. Akkermans[a], E. Noriega Fernandez[a], F. Logist[a], J. Van Impe[a,*]

[a]*KU Leuven, Department of Chemical Engineering, BioTeC+ & OPTEC, Gebroeders De Smetstraat 1, 9000 Ghent, Belgium*

**Abstract**

In predictive microbiology, the (induced) lag-phase is a phenomenon of specific interest, as it has a large impact on the assessment of safety and quality of food products. This lag phase has been studied mostly on a macroscopic level. However, a quest for more mechanistically-based predictive models has started, for example through the integration of a metabolic reaction network into widely used macroscopic model structures. This multi-scale modeling approach is called dynamic metabolic flux analysis (dMFA). In this contribution, a recently developed algorithm for dMFA is used to estimate the metabolic fluxes in *Escherichia coli* K12 during an experimentally induced lag phase through a sudden shift in temperature. To study this phenomenon, controlled bioreactor experiments were performed: on the one hand at a fixed and optimal temperature for growth (37°C), and on the other hand starting at 20°C, with a sudden temperature shift to 37°C during the exponential growth, inducing an intermediate lag phase. The evolution of biomass and metabolite concentrations was monitored during these experiments. After dMFA analysis of the gathered measurements, some interesting patterns in metabolic activity during the different growth phases are revealed. The described case study is a first practical test case to assess the capabilities of the recently developed dMFA methodology in an experimental predictive microbiology setting.

*Keywords:* Dynamic metabolic flux analysis, B-spline parametrization, Experimental validation, Predictive microbiology

## 1. Introduction

The concept of predictive microbiology entails that a detailed knowledge of the behavior, i.e., growth, survival and inactivation, of microorganisms in food products can be expressed in mathematical models,

*Corresponding author
*Email address:* jan.vanimpe@kuleuven.be (J. Van Impe)

which enable an objective evaluation of the microbiological safety and quality of foods [22]. An important phenomenon studied in predictive microbiology is the *lag phase*, which is a period in which no growth occurs as microorganisms have to adapt to a (sudden) change in environmental conditions. This can happen because of inoculation of the organism in a new medium, in which case the phase is called *initial lag*, or due to a change in one or more environmental variables, e.g., temperature or pH, during an exponential growth phase, in which case the period is called an *intermediate* or *induced lag phase*. Because of the frequent changes in environmental conditions taking place during the production, distribution and consumption of food products, a good understanding of the lag phase is of vital importance for the assessment of microbial safety and quality of food products. Once the influence of environmental conditions on the occurrence and the length of the lag phase is determined, shelf-life of food products can be determined more accurately, and strategies can be developed to inhibit the growth of microorganisms by keeping them in this lag phase for extended periods of time.

For these reasons, the lag phase has been studied extensively in recent years in predictive food microbiology [27]. Examples include studies of (*i*) the initial lag phase at different temperatures [4, 35], or for individual cells [16]; (*ii*) the lag phase due to a sudden shift in temperature for different microorganisms, e.g., *Escherichia coli* [28], *Listeria monocytogenes* [12] and *Klebsiella oxytoca* [23]; and (*iii*) the lag phase due to a fluctuating temperature and water activity [25]. The main difference between [27] and this paper is the focus on, on the one hand the macroscopic lag phase description (the past work), and on the other hand the estimation of intracellular fluxes during this lag phase (the present work). In the former, the lag phase is quantified on a macroscopic level so pH and dissolved oxygen are controlled on-line while the off-line measurements concerned only the cell density, glucose and acetate concentration. In contrast, in order to estimate the intracellular behavior, a more detailed set of measurements are taken, i.e., biomass, $O_2$, $CO_2$, ammonium and hydrocarbons. A more detailed description is provided in the materials and methods section.

All these studies focus on a description of the lag phase from a macroscopic point of view, i.e., the models describing the phenomenon are based on the cell and population level. However, a quest for more mechanistically-based predictive models has started recently [21]. The applicability and reliability of existing models will be improved by looking inside the black box and unraveling the underlying mechanisms [9]. One way of adding mechanistic information to a macroscopic model is by integrating the information contained in a *metabolic reaction network* into the model [29]. These networks are a representation of (a part of) the



Journal homepage: https://www.journals.elsevier.com/food-and-bioproducts-processing/
Original file available at: http://www.sciencedirect.com/science/article/pii/S0960308516301110

metabolic reactions occurring inside the cell, and between the cell and its environment. The rates of these reactions, the so-called *fluxes*, provide important information on the distribution of resources by the cell over the different pathways in its metabolism. These fluxes can be calculated based on extracellular concentrations and fluxes through *(dynamic) metabolic flux analysis* techniques [36, 2]. Furthermore, metabolic reaction networks can be used to build predictive models through *(dynamic) flux balance analysis*, in which a cellular objective is used to obtain the fluxes [19]. One application of this technique in predictive microbiology is given in Metris et al. [24], in which osmotic stress is modeled through flux balance analysis. Other potential application is in the assistance of drug development to determine relevant/critical pathways [5]. Furthermore, bioreactor operation has been shown to increase when the metabolic network is considered in the optimization procedure [11].

The application of metabolic reaction network-based techniques in predictive microbiology is, however, still in the starting blocks. To prove the relevance of this methodology in this field, more practical applications are needed. The aim of this contribution is the experimental validation of a recently developed technique for dynamic metabolic flux analysis [30, 31] to estimate the dynamic metabolic fluxes in *E. coli* K12 during a lag phase induced through a sudden shift in temperature. The developed method enables to take the nonlinear nature of the dMFA-problem into account. The free fluxes are parametrized with B-splines to ensure smoothness of the flux profiles and state-of-the-art tools for solving dynamic optimization problems are used to solve the dMFA-problem as a dynamic parameter estimation problem. The controls and states are discretized by an orthogonal collocation scheme on finite elements and as solver an interior point optimizer, i.e., IPOPT [34], is used. For the calculation of exact first and second order derivatives, CasADi [1] provides the automatic differentiation. The experimental materials and methods used in the bioreactor experiments are described in Section 2, along with the employed metabolic reaction network. Section 3 starts with a description of the measurements gathered during these experiments, subsequently the employed dynamic flux estimation algorithm is briefly described. The section concludes with a detailed description of the results from the dynamic metabolic flux estimation. Finally, Section 4 summarizes the main conclusions of this contribution.

## 2. Materials and methods

### *2.1. Culture media*

Two different culture media were used for the experiments: (*i*) a liquid medium for cell reactivation, preculture and growth experiments in the bioreactors, and (*ii*) a solid culture medium to quantify cell

concentration by viable plate counting.

#### *2.1.1. Liquid medium*

This medium is the M9 minimal medium described for *E. coli* K12 in [15] and [26]. It is based on glucose as a limiting carbon source, a phosphate buffer, salts and micronutrients. This medium contains, per liter: 5 g D-glucose (Acros), 1 g $NH_4Cl$ (VWR Prolabo), 0.5 g NaCl (VWR Prolabo), 6 g $Na_2HPO_4$ anhydrous (Acros), 3 g $KH_2PO_4$ (Fisher Scientific). The following components are then added separately from a stock solution: 2 mL of 1 M $MgSO_4 \cdot 7\ H_2O$ (Chem-Lab), 1 mL of 0.1 M $CaCl_2$ (Sigma), 0.3 mL of 1 mM thiamine HCl (Sigma) and 10 mL of trace element solution. This trace element solution contains, per liter: 1 g $FeCl_3 \cdot 6\ H_2O$ (Acros), 0.18 g $ZnSO_4 \cdot 7\ H_2O$ (Chem-Lab), 0.12 g $CuCl_2 \cdot 2\ H_2O$ (Fluka), 0.12 g $MnSO_4 \cdot H_2O$ (Acros) and 0.18 g of $CoCl_2 \cdot 6\ H_2O$ (Riedel de Haën). This growth medium is sterilized according to the protocol in Nanchen et al. [26]. The different solutions containing glucose, salts and phosphate buffer are separately autoclaved (VWR Vapour Line) and mixed afterwards. Then, the separately sterilized micronutrient solution is added after cooling of this mixture. Thiamine HCl is filter sterilized through a 0.2 $\mu$m pore size filter (TRP Filtermax) due to thermal instability, as well as the trace element solution, to prevent precipitation of any poorly soluble salts during or after autoclaving.

#### *2.1.2. Solid medium*

The plate medium (BHIA) contains 37 g/L of Brain Heart Infusion (BHI, Oxoid Ltd.) and 12 g/L of Agar Technical (Oxoid Ltd.).

### *2.2. Experimental protocols*

The experimental protocol for the batch bioreactor experiments starts with cell reactivation, preculture and inoculum preparation. Subsequently, the autoclaved bioreactor and growth medium are inoculated. Measurements of viable cell concentration and metabolites are performed on samples taken at regular time intervals.

#### *2.2.1. Preculture*

The bacterial strain used for all experiments is *E. coli* K12 MG1655. The strain is acquired from the Coli Genetic Stock Center (Yale University), and stored at -80°C in a mixture of BHI broth and 20% (v/v) of glycerol.

For all the experiments, a purity plate is prepared in BHIA by streaking a loopful of the frozen stock culture, and incubated at 37°C (Binder). After 24 h, a single colony is inoculated in a 50 mL Erlenmeyer flasks containing 20 mL of the liquid growth medium previously defined. After 24 h incubation at 37°C, 80 $\mu$L are transfered to six 250 mL Erlenmeyer flasks containing 80 mL of fresh growth medium and incubated under the same conditions. From the second preculture, 45 mL is transferred into six 50 mL sterile tubes with cap (Falcon). The tubes are centrifuged at 30°C and 5000 rpm for 15 minutes (Eppendorf Centrifuge 5810R), the supernatant is discarded and the respective pellets are resuspended in 5 mL fresh growth medium. Finally, the bioreactor is inoculated with these cell suspensions (30 mL) to obtain an optical density of 0.03.

### *2.2.2. Bioreactor*

The batch experiments are conducted in a 5 L bioreactor (BioFlo 3000, New Brunswick Scientific). The reactor is filled with 3.5 L of liquid growth medium. It is equipped with online sensors and PID control for temperature, pH, aeration, agitation speed and dissolved oxygen (DO). The control of all variables is executed by the AFS BioCommand software (Advanced Fermentation Software, New Brunswick Scientific). An antifoaming agent (Y-30, Sigma) is added to prevent foam formation on the medium.

pH is measured using a glass electrode (Mettler Toledo), which is calibrated before autoclaving the reactor. During the experiment the pH is kept constant at a value of 7 for all experiments, by adding sterile acid ($H_2SO_4$) or basic (KOH) solutions using peristaltic pumps. The total amount of acid and base added is recorded by the software. Aeration is achieved by an air sparger system, pumping sterile air into the reactor at a constant rate of 2 L/min. Mixing is done by a six bladed turbine rotating at 400 rpm. Baffles are introduced into the reactor to prevent vortices and to create additional turbulence. The temperature is measured by a resistance temperature probe. Temperature is controlled through the external jacket of the reactor.

Two bioreactor experiments were performed. The first one was conducted at the optimal temperature of 37°C, the second one was started at 20°C, after which, during the exponential growth phase, a sudden temperature shift to 37°C was implemented to induce an intermediate lag phase. This specific temperature shift was chosen because it leads to a relatively long and pronounced lag phase, as described in Swinnen et al. [28].

### *2.2.3. Measurements*

Samples were taken at regular intervals via the sampling port of the bioreactor. The volume of the sample was always approximately 12 mL, and was measured to keep track of the reactor volume. For the

first experiment, measurements were started immediately after inoculation, and were taken once every hour during 11 h. At the end of this period, the cell population was just approaching the stationary phase. In the second experiment, measurements were started during the exponential growth phase, 12 h after inoculation, i.e., three hours before the temperature shift, and continued for 12 h. Samples were taken every 30 min in the three hours before and three hours after the temperature shift, and every hour for the rest of the experimental horizon, resulting in a total of 19 samples.

*Biomass.* Dry cell weight and viable plate counting methods are used in this work to measure biomass concentration.

For each sample, serial decimal dilutions were prepared and plated in duplicate onto BHIA using a spiral plater (Eddy Jet, IUL Instruments). The plates were incubated at 37°C for 24 h. The remaining sample was filtered through a 0.2 $\mu$m filter (Filtropur S0.2) and stored in the freezer (Siemens) at -20°C for later measurements.

The dry weight is measured by filtering the sample and weighing it after drying. A filter of 0.2 $\mu$m pore size (Porafil membrane filter) is dried for 24 h in a 105°C oven (Binder). The filter is then placed in a desiccator for 15 min and weighed. Subsequently, a 5 mL sample is filtered and the filter is allowed to dry for 24 h in a 105°C oven. Then, the filter is placed in a desiccator and weighed again. The difference between the final weight and the initial weight is the dry weight of the sample. This sample dry weight includes the cells and any retained particles from the medium.

A calibration curve was determined to convert viable plate counts into grams of dry weight (gDW). The viable plate counts and dry weight were determined for cell suspensions with different concentrations. The following equation is derived for the conversion of CFU/mL to gDW/mL:

$$DW = 6.07 \times 10^{-13} \cdot CFU \tag{1}$$

with $DW$ the cell dry weight [gDW/mL] and $CFU$ the total viable plate count [CFU/mL]. This equation is used to prepare the biomass data for usage in the dMFA algorithm.

*Hydrocarbons.* Most of the extracellular metabolites are hydrocarbons, either sugars such as glucose or acids such as acetic acid. These hydrocarbon concentrations were measured via HPLC. For each filtered sample

from the bioreactor, dilutions were prepared with ultrapure water and analyzed using an Agilent 1200 series HPLC equipped with absorbance and refractive index detectors and a fermentation monitoring column (Bio-Rad). The analysis was performed at 50°C using 5 mM $H_2SO_4$ as mobile phase. The flow rate was 0.6 mL/min and the injection volume was 20 $\mu$L. Glucose was detected by the refractive index detector (RID) and organic acids are detected by the UV-VIS absorbance detector at 210 nm.

*$O_2$ and $CO_2$.* The $O_2$ and $CO_2$ headspace concentrations are measured using off-gas analysis. The outlet gas is sent through a condenser to a gas analysis unit (BlueSens). The incoming air via the sparger system is assumed to be known and calibrated at 20.97% $O_2$ and 0.04% $CO_2$. The Oxygen Uptake Rate (OUR) [mol/L h] and $CO_2$ Transfer Rate (CTR) [mol/L h] can be calculated via a mass balance. The following formulas apply for the respective rates [17]:

$$OUR = \frac{F^{in}}{V(t) \cdot V_m} \cdot \left( x^{in}_{O_2}(t) - x^{out}_{O_2}(t) \cdot \frac{1 - x^{in}_{O_2}(t) - x^{in}_{CO_2}(t)}{1 - x^{out}_{O_2}(t) - x^{out}_{CO_2}(t)} \right) \tag{2}$$

$$CTR = \frac{F^{in}}{V(t) \cdot V_m} \cdot \left( x^{out}_{CO_2}(t) \cdot \frac{1 - x^{in}_{O_2}(t) - x^{in}_{CO_2}(t)}{1 - x^{out}_{O_2}(t) - x^{out}_{CO_2}(t)} - x^{in}_{CO_2}(t) \right) \tag{3}$$

with $F_{in}$ the incoming gas flow rate [L/h], $V(t)$ the reactor volume [L], $V_m$ the standard molar volume [L/mol], $x^{in}_{O_2}(t)$ and $x^{in}_{CO_2}(t)$ the molar fraction of $O_2$ and $CO_2$ in the incoming air, and $x^{out}_{O_2}(t)$ and $x^{out}_{CO_2}(t)$ the molar fraction of $O_2$ and $CO_2$ in the outgoing air.

*Ammonium.* The ammonium concentration is spectrophotometrically determined using a predosed cuvette test (LCK303, Hach Lange) based on the indophenol blue method. The use is facilitated by a bar code on the cuvette, allowing automatic detection and simple reading of the measurement on the spectrophotometer (Hach Lange DR5000).

### *2.3. Metabolic reaction network*

The metabolic reaction network that is used for the dMFA estimation, is an adapted version of the network found in Antoniewicz et al. [3]. The full list of reactions in the metabolic reaction network used in the estimations can be found in the Supplementary data.

The network consists of the most important parts of the *E. coli* metabolism, including the central carbon metabolism and amino acid biosynthesis. The adaptations made to the model of Antoniewicz et al. [3] include the removal of a number of pathways. First, the Entner-Doudoroff pathway is removed, as its ATP

yield is lower than glycolysis and therefore considered to only be active when glycolysis is not possible (due to lack of enzymes). The 1,3-propanediol biosynthesis pathway is removed as this pathway is only present in metabolically engineered strains. Furthermore, the amphibolic reactions, which are only present in gluconeogenesis, are removed. Phosphoenolpyruvate carboxylation is considered, as it provides an alternative route into the tricarboxylic acid cycle, which is important in this study. Additionally, the transhydrogenation of NADH is removed as this balance is already provided by another reaction. The exchange fluxes are chosen in function of the metabolites that were identified experimentally. In this case, alpha-ketoglutarate and acetate were exchanged in addition to the glucose uptake. Citrate was not exchanged and therefore its uptake is excluded from the network. These modifications result in a network consisting of 58 metabolites and 62 reactions. If pseudo steady-state is assumed, this means 4 free fluxes have to be determined. A mathematical network description can be considered as a linear system of equations which needs to be solved. So, the notion of 4 free fluxes means that once values for these free fluxes are established, e.g., on the basis of experiments, the remaining 58 fluxes are implicitly set by the reaction network description (the linear model network description). If these are subsequently determined during the estimation algorithm based on the measurements taken in the experiment, the remaining fluxes are also set. However, any uncertainty in this estimation propagates through to all fluxes.

The network describes the central carbon metabolism of *E. coli*. This comprises the fueling reactions which provide the cell with energy and precursors for building blocks. It starts with a nutrient or carbon source such as glucose which is broken down to pyruvate through glycolysis. Pyruvate is then further converted via either aerobic respiration or fermentation, resulting in a number of different possible end products. The most important pathways are glycolysis, the pentose phosphate pathway (PP) and the tricarboxylic acid cycle (TCA). Figure 1 shows these pathways and their sources and products graphically. As a simplification, not all intermediate metabolites are shown.

#### *2.3.1. Glycolysis*

The set of reactions participating in the breakdown of glucose to pyruvate is called glycolysis. The extracellular carbon source, for example glucose, enters the glycolysis pathways by phosphorylation at the C-6 position, using one ATP molecule. The glycolysis pathway is the main energy conversion pathway in the breakdown of sugars to pyruvate. It results in a net gain of 2 ATP molecules and 2 NADH molecules per molecule of glucose. Glucose enters the pathway through phosphorylation to glucose-6-phosphate (G6P) using 1 ATP. It is then further broken down to fructose-6-phosphate (F6P) and again phosphorylated to 1,6-fructose-biphosphate (FBP) using another ATP. This FBP is splited into two C3 molecules which are

interconvertible, glyceraldehyde-3-phosphate (G3P) and dihydroxyacetone-phosphate (DHAP) and further converted to pyruvate. In this process 4 molecules of ATP are released, resulting in the previously mentioned net gain of 2 ATP molecules.

### *2.3.2. Pentose phosphate pathway*

The PP pathway is an oxidative pathway producing necessary precursor metabolites for biosynthesis routes. Additionally, it is a source of NADPH, which provides reducing power in biosynthesis. It starts from G6P, the same starting point as the glycolysis pathway. This is converted to pentose-5-phosphate (P5P), which results in 2 molecules of NADPH. This P5P can then be further converted to F6P or G3P through a non-oxidative pathway, which links it back to the glycolysis pathway. The ratio at which glucose is broken down through the PP pathway depends on the precursor, ATP and NADPH requirements of the cell. Usually these are increased with biomass production.

### *2.3.3. Tricarboxylic acid cycle*

The tricarboxylic acid cycle or Krebs cycle is responsible for the complete oxidation of pyruvate in aerobic circumstances. First, pyruvate is decarboxylated to Acetyl-Coenzyme A (AcCoA) and $CO_2$. AcCoA enters the TCA cycle by condensation with oxaloacetate (OAA) to citrate. The cycle is then completed through a number of intermediates. In this cycle $NAD^+$ and FAD are used as oxidizing agents. The cycle gives a net result of 3 $CO_2$, 1 ATP, 4 NADH and 1 $FADH_2$ molecules. The cycle can then only continue if NADH and $FADH_2$ are recycled to $NAD^+$ and FAD. This occurs through oxidative phosphorylation which results in a gain of 3 ATP per NADH molecule and 2 ATP per $FADH_2$ molecule. The TCA cycle is responsible for a net gain of 30 ATP molecules per molecule of glucose that is metabolized.

The TCA cycle is of great importance in many microorganisms. It provides a large portion of the cell's energy demand and its intermediates, such as alpha-ketoglutarate (AKG), are important precursors for biosynthesis of amino acids. Some intermediates can be secreted, among which AKG and acetate. If oxygen is not present, the cell cannot use the TCA cycle as oxidative phosphorylation becomes impossible. In this case, pyruvate is metabolized through fermentation, leading to production of a number of carbohydrates such as ethanol, lactic acid and formic acid.



Journal homepage: https://www.journals.elsevier.com/food-and-bioproducts-processing/
Original file available at: http://www.sciencedirect.com/science/article/pii/S0960308516301110

Postprint version of paper published in Food and Bioproducts Processing 2017, vol. 102, p. 1-19.
**The content is identical to the published paper, but without the final typesetting by the publisher.**

## 3. Results and Discussion

### *3.1. Batch growth experiment measurements*

To study the induced lag phase of *E. coli* K12 using dMFA, two bioreactor experiments were performed to collect data for the dMFA estimation. The first one was performed at the optimal growth temperature for *E. coli* of 37°C, while the second experiment included a sudden temperature shift to induce an intermediate lag phase. The temperature profile for this experiment is shown in Figure 3. The measurements gathered during these experiments are now described.

#### *3.1.1. Biomass*

The growth curves for both experiments are given in Figure 3. In the experiment at the optimal growth temperature, after a short initial lag phase of about one hour, the cells start to grow exponentially until a slight decrease in the growth rate is visible at the end of the experiment, indicating the start of the stationary phase. The data from the optimal temperature experiment are, however, only used as an initial proof of concept for the dMFA methodology, so no extensive conclusions need to be drawn from this data. The measurements of the second experiment start in the exponential growth phase at low temperature, which is followed by a short intermediate lag phase, a second exponential phase at optimal temperature and a clear stationary phase at the end of the experiment.

#### *3.1.2. Oxygen and carbon dioxide headspace concentrations*

The oxygen and carbon dioxide concentrations in the headspace of the reactor are shown in Figure 4 for both experiments. As these measurements are performed online, a far higher sampling rate was obtained than for the other measurements.

In the optimal temperature experiment, the oxygen concentration decreases until about 10 h, while the carbon dioxide concentration increases accordingly over the same time range. This is the time during which glucose is metabolized and aerobic respiration takes place due to activity of the TCA cycle. When glucose is depleted, the uptake of oxygen slows down and finally the growth stops completely. In this phase, acetate is still metabolized, but this leads to a lower activity of the TCA cycle. Also in the temperature shift experiment, this behavior can be observed. The oxygen concentration decreases initially, during the first exponential growth phase, but starts increasing again when the temperature shift is executed. Three effects contribute to this change. Due to the temperature change, the oxygen saturation concentration in the medium decreases, but the oxygen transfer coefficient increases, i.e., although oxygen transfer to the medium



Journal homepage: https://www.journals.elsevier.com/food-and-bioproducts-processing/
Original file available at: http://www.sciencedirect.com/science/article/pii/S0960308516301110

will be faster because of the higher transfer coefficient, less oxygen can be potentially transferred because of the lower saturation level. Also, due to the slower growth because of the temperature shift and induced lag phase, an increase in the oxygen level can be expected. In the carbon dioxide level, there is no clear effect from the temperature shift. After the temperature shift, the same phenomena as in the experiment at the optimal temperature for growth can be observed.

#### *3.1.3. Metabolite concentrations*

The measurements for the concentrations of glucose, acetate, AKG and ammonium are summarized in Figure 5 for both experiments. Glucose and ammonium are taken up by the cells until depletion over the course of the experiment. Initially the uptake rates are slow, but as biomass growth reaches the exponential phase, the glucose and ammonium uptake rates increase as well. Parallel to the increase in biomass and the uptake of glucose, acetate and AKG are secreted. At the end of the exponential growth phase, after the depletion of glucose, acetate is taken up again in the temperature shift experiment.

### *3.2. dMFA flux estimation*

Finally, after gathering the necessary data and choosing a metabolic reaction network, the dMFA flux estimation can be performed. In this section, the preparation of the data for flux estimation is described, along with the final dMFA problem to be solved and the results, i.e., the estimated fluxes.

#### *3.2.1. Data preparation*

The stoichiometry in the network is only valid for molar units, therefore all concentrations gathered must be converted to either mmol/L or mCmol/L, for the biomass. The plate counts are converted to dry weight using the calibration curve previously defined in Equation (1). An empirical formula for the molar composition of *E. coli* is found in [33] as $C_4H_7O_2N$. This suggests that a conversion factor of $\frac{4}{101}$ can be used to convert gDW to Cmol. Dividing by the reactor volume results in the biomass concentration.

In practice, only one measurement at each sampling point and for each measured variable has been performed. Measurement variances are based on qualitative arguments for the measuring techniques and are assumed to be Gaussian distributed with zero mean. So, the general variability in the time series data is considered as follows:

- For the AKG, acetic acid and biomass measurements: the standard deviation was set at 5% relative to the measured value.
- For glucose and ammonia: the standard deviation was set to 20% relative to the measured value.
- All non-offgas variances were subject to a minimum value of 0.001, so if a calculated variance based on the above considerations was smaller than 0.001, it was set to 0.001 exactly.
- For oxygen a standard deviation of 0.04 is assumed.
- For carbon dioxide a standard deviation of 0.015 was taken.

*Remark*: Since the offgas data are sampled at a much higher frequency than the offline measurements, these data are downsampled so that they were in line with the offline measurements, i.e., only the data at the time points where offline samples were taken, were retained.

Finally, both datasets have been visually inspected for outliers. Based on the biological insight for the two experiments that both glucose and ammonium are consumed at a very slow rate initially, three points have been identified as outliers in the optimal temperature experiment: the glucose measurements at 1 h and 5 h, and the ammonium measurement at 5 h. These points have been removed from the dataset. Also in the temperature shift experiment, three outliers are detected. The resulting sets of measurements for both experiments after these modifications are shown in Figures 6 and 7. For the optimal temperature experiment, 7 variables were measured at 12 time points, after removing 3 outliers, this results in a dataset of 81 data points with associated variances. For the temperature shift experiment, the dataset contains measurements for 7 variables at 19 time points, with 3 outliers, this results in 130 data points with associated variances.

### *3.2.2. The dMFA estimation problem*

Based on the specified network and the measurements that were gathered, the following dMFA problem is defined:

$$\underset{\hat{\mathbf{u}}(t),\mathbf{x}_0,\mathbf{K}}{\text{minimize}} \sum_{i=0}^{n_{\text{time}}} \sum_{j=1}^{n_{\text{out}}} \left( \frac{y_j(t_i) - m_{ij}}{\sigma_{ij}} \right)^2 \tag{4}$$

subject to:

$$\dot{\mathbf{x}}(t) = \mathbf{S}_\mathrm{e} \cdot \mathbf{K} \cdot \hat{\mathbf{u}}(t) \cdot \mathbf{q}_\mathrm{bio}^\mathsf{T} \cdot \mathbf{x}(t) \tag{5}$$

$$0 = \frac{\mathbf{x}_\mathrm{head,in} \cdot F \cdot M}{V_\mathrm{liq}} - \frac{\mathbf{x}_\mathrm{head}(t) \cdot F \cdot M}{V_\mathrm{liq}} \cdot \frac{1 - \sum_i \mathbf{x}_{\mathrm{head,in},i}}{1 - \sum_i \mathbf{x}_{\mathrm{head},i}(t)} - \mathbf{K}_\mathrm{gas} \cdot \hat{\mathbf{u}}(t) \cdot \mathbf{q}_\mathrm{bio}^\mathsf{T} \cdot \mathbf{x}(t) \tag{6}$$

$$\mathbf{x}(0) = \mathbf{x}_0 \tag{7}$$

$$\mathbf{y}(t) = \begin{bmatrix} \mathbf{x}_\mathrm{sel} & \mathbf{x}_\mathrm{head} \end{bmatrix}^\mathsf{T} \tag{8}$$

$$\mathbf{z}(t) = \mathbf{I}_\mathrm{irr} \cdot \mathbf{K} \cdot \hat{\mathbf{u}}(t) \tag{9}$$

$$\mathbf{x}(t) \geq 0, \mathbf{x}_\mathrm{head}(t) \geq 0, \mathbf{z}(t) \geq 0 \tag{10}$$

This is an infinite dimensional estimation problem using a weighted least squares objective function in which the difference between the simulated output and the measurements is minimized. In this objective function, $n_\mathrm{time}$ is the total number of time points at which measurements are taken, $n_\mathrm{out}$ is the number of outputs of the system, i.e., the total number of variables that are measured at each time point, $y_j(t_i)$ is the simulated output of the model for output $j$ at time $t_i$, and $m_{ij}$ and $\sigma_{ij}$ are the measurement and the standard deviation, respectively, for output $j$ at time $t_i$.

The first two constraints define the dMFA model. In these equations, $\mathbf{x}(t)$ is the vector of metabolite concentrations, while $\mathbf{x}_\mathrm{head}(t)$ contains the headspace molar fractions of oxygen and carbon dioxide, as defined in Table 1. The vector $\mathbf{x}_0$ denotes the initial condition of the system of ordinary differential equations which is also estimated in the discussed algorithm (and common practice in (moving horizon) estimation [18, 32]). In Equation (5), $\mathbf{S}_\mathrm{e}$ contains the rows of the stoichiometric matrix corresponding to the medium metabolites and the biomass, $\mathbf{K}$ is the null space basis of the intracellular stoichiometric matrix, $\hat{\mathbf{u}}(t)$ are the free fluxes to be estimated, and $\mathbf{q}_\mathrm{bio}^\mathsf{T}$ is a row vector that selects the biomass from $\mathbf{x}(t)$. The full description of the dynamic model is given in Vercammen et al. [30] and Vercammen et al. [31].

Apart from the dynamic model, algebraic equations (Equation (6)) that describe the evolution of the headspace states, based on Equations (2) and (3), are added. In this equation, $\mathbf{K}_\mathrm{gas}$ is the matrix composed of the rows of $\mathbf{K}$ corresponding to the oxygen and carbon dioxide exchange fluxes, i.e., reactions 60

and 61 according to the list of reactions in the Supplementary data. The fractions in the inlet gas $\mathbf{x}_{\text{head,in}}(t)$ are defined in Table 1, as well as the metabolite states that were selected for measurement $\mathbf{x}_{\text{sel}}$. Apart from these selected metabolite concentrations, also the headspace mole fractions are measured. Irreversibility constraints on the fluxes are represented in Equation (9). In this equation, the values in $\mathbf{z}(t)$ represent the irreversible fluxes, and these values are required to be positive. These values are selected from the full set of fluxes by the matrix $\mathbf{I}_{\text{irr}}$. Descriptions and numerical values for the parameters $F$, $M$ and $V_{\text{liq}}$ are given in Table 2.

***Solution strategy*** In this contribution, the dMFA methodology described previously in Vercammen et al. [30] and Vercammen et al. [31] is employed. The infinite-dimensional estimation problem ((4)-(10)) is transformed into a series of nonlinear programming problems using direct collocation on finite elements for the states and B-spline parametrizations for the fluxes. In the developed approach, every free flux is parametrized as a polynomial function, based on B-spline basis functions [14]. These B-spline functions are defined by three factors: (i) the degree $k$, (ii) the location of the $g+2$ knots $t_0, t_1, \ldots, t_g, t_{n_{\text{time}}}$, of which the middle ones are the $g$ internal knots, and (iii) the control points, or spline parameters $p_u$. The start and end knots ($t_0$ and $t_{n_{\text{time}}}$) are fixed at the start and end times of the experimental time horizon. Consequently three entities are left to be calibrated: the number of internal knots, their locations, and the spline parameters. The Cox-de Boor algorithm [13] allows for an efficient evaluations of these spline functions. Each free flux is represented by a corresponding spline function, and thus has its own set of internal knot locations and spline parameters. The total number of parameters ($n_{\text{p}}$) for each free flux is given by:

$$n_{\text{p}} = 2 \cdot g + k + 1 \; . \tag{11}$$

Three aspects with respect to splines need to be considered. The first is to constrain a knot to its specific measurement interval in order to prevent knots from straying too far from their initial optimal location. The second is inserting a knot at the end of a specified time frame. Knot insertion, a feature inherent to B-splines, is used to insert a knot without changing the spline profile. This way, the next optimization can be started from a good initial guess, with an extra knot inserted. This also accounts for the bounds which were placed on previously added knots, i.e., the knot is inserted in the time frame at the end, where no knot has yet been inserted. The last aspect is the spline prolongation, which only changes the ending knot $t_{n_{\text{time}}}$ to a new value. This means that the spline profile is slightly changed, because the spline parameters stay the same, but is still close to the previous profile.

The dynamic optimization problem needs to be discretized to solve it. In this work, as mentioned previously direct collocation on finite elements is chosen [8]. The resulting nonlinear programming problems are solved using the interior-point optimization routine IPOPT [34]. Gradient, Jacobians and Hessian are generated exactly using automatic differentiation with CasADi [1]. Note that in [20] a different type of B-spline dMFA approach is presented which is only linear in the parameters to be optimized. It requires the predetermined positioning of knots and does not consider the null space as an additional optimization variable. It is computationally cheaper but can result in a less approprate fit due to the lack of decision variables.

*Incremental free flux estimation algorithm* The number of internal knots directly controls the number of other parameters. As the least squares objective decreases with increasing number of parameters, the optimal number of knots tends to infinity, rendering a perfect fit of the measurement noise instead of the trend. For this reason, a mechanism to prevent overfitting has to be used. Furthermore, although the polynomial spline functions are linear functions of the spline parameters, both the system of ordinary differential equations is nonlinear, and the splines are also nonlinear in the knot locations. These nonlinearities can result in local minima in the optimization solution. To address these issues, a systematic, incremental strategy for estimating the free flux parameters and knot locations has been devised, based on Akaike' s model discrimination criterion. This criterion [10] is used to discriminate between different models which describe the same phenomenon. It takes into account both the model error and the number of parameters needed. The corrected AIC criterion ($\mathrm{AIC}_c$) is used, as it is more suited in cases where the number of measurements is close to the number of parameters:

$$\mathrm{AIC}_c = f + (2 \cdot n_\mathrm{p} \cdot (n_\mathrm{p} + 1))/(n_\mathrm{meas} - n_\mathrm{p} - 1) \;, \tag{12}$$

with $f$ the weighted least squares error, as defined in Equation (4), and $n_\mathrm{meas}$ the total number of measurements. From this definition it is clear that $n_\mathrm{meas} \geq n_\mathrm{p} + 2$, as otherwise the denominator can become zero or negative.

The method is started by estimating splines without knots on a reduced dataset (Step 1), i.e., the first $l$ time points, where $l$ is the number of time points needed to make sure that the denominator of Equation (12) is strictly positive. After selecting the correct number of time points, the polynomials are fitted and the optimal AIC value is saved as $F$, together with the optimal splines $U$ (Step 2). After this, additional optimization

problems are generated, in which every time one knot is inserted into one free flux spline at a time (i.e., in Step 3a as long as there is a sufficient number of measurements to add a new knot, else a time point is added and the splines are prolongated in Step 3b if there are still points to be added). The problems are solved, and the minimum AIC value over these problems is saved as $AIC^*$, along with the corresponding optimal splines $U^*$, in which there is now one knot in one of the splines (Step 4a). Now, two possibilities arise. If $AIC^*$ is smaller than $F$, a new, better solution is found (Step 5a), and another knot can be added. If $AIC^*$ is bigger than $F$, however, the old optimum was better than the new one, and so the old one is kept. An optimal solution for this number of time points has been found, and a new time point can be added if there is still one left to be added. After adding the time point, the splines are prolonged, and the prolonged problem is solved to get the new starting values for $F$ and $U$ for the next iteration. Once all time points have been added, $U$ contains the final, optimal set of free flux profiles for the specified dataset. For a more detailed explanation, the interested reader is referred to [30]. The overall approach employed in the incremental free flux algorithm is schematically depicted in Figure 2.

### *3.2.3. Results for the optimal temperature experiment*

In this section the flux estimation for the experiment at optimal temperature is described. The full dMFA methodology is performed on the data for the optimal temperature experiment, including the estimation of the null space basis for the intracellular stoichiometric matrix. This resulted in spline knot locations located at 8.26 and 9.75 h for the first free flux, 6.50 h for the second, 4.35 h for the third and 5.67 h for the last free flux. Since the null space basis is estimated, these free fluxes do not correspond to real fluxes, so these knot locations are not that informative, but are mentioned here for completeness. The estimated outputs and fluxes will be discussed below in more detail.

*Goodness of fit with respect to measurements.* First, the calculated concentration profiles and exchange fluxes resulting from the optimal flux distribution are compared to the measurements of these concentrations. A good fit ensures that the calculated flux distribution corresponds to the real situation. If no good fit with the measurements is observed, either the metabolic network is not sufficient to describe the intracellular dynamics or the measurements are not accurate enough. In the former case any reduction of the complexity causes inaccuracy in the model, for example the potential absence of reactions that take place due to the choice of network model. In the latter case, the measurements are too inaccurate to allow an appropriate fit

in correspondence with the model.

The fitted measurement data is depicted in Figure 6. If the estimated concentration profiles are within the confidence bounds of the measurements defined by the error bars (note that these denote the 95% confidence interval), the fit is considered to be satisfactory. In general, these figures show a good fit of the data, with a slight overestimation of glucose, and a slight underestimation of the last point in the AKG, acetate and biomass fits. However, in general, the fit is satisfactory. When the reduced chi-squared measure is computed, a value of 3.36 is obtained. The optimal value for this measure is equal to 1. For this experiment, it indicates that not all data is completely captured which is in line with the observations of e.g., the glucose evolution where all residuals have the same sign.

*Analysis of metabolic pathways.* Based on the estimated fluxes, the activity of the important pathways in the central carbon metabolism is now analyzed. This serves as a reference for optimal growth, which can later be compared to results for the induced lag phase experiment. The differences in metabolic activity during this phase are then highlighted.

The profiles of the fluxes are divided into four figures according to the pathway which they belong to: the glycolysis pathway, the PP pathway, the TCA cycle or the exchange fluxes. Figure 8 displays the reactions of glycolysis and pyruvate decarboxylation. This is the direct route from glucose uptake to the entry into the TCA cycle. This cycle cannot take place without the necessary regeneration of strong reductants provided by this oxidative phosphorylation. For this reason they are shown in Figure 10 together with the oxidative phosphorylation reactions. These reaction pathways are active when the cell requires energy, as they generate the most ATP. The PP pathway, shown in Figure 12 is an alternative to glycolysis providing the cell with precursors for biosynthesis of macromolecules. Finally, the exchange fluxes are shown in Figure 14. From these figures it is possible to link activity of certain pathways to phases in the growth of the culture.

First, the glycolysis and TCA cycle pathways become active, about 2 h after inoculation. The cells require energy in the form of ATP in order to grow, which is provided in abundance by the TCA cycle. Due to imbalance of the partial reactions in these pathways, secretion of accumulating intermediates occurs, mainly in the form of acetate, as demonstrated by the measurements and the calculated exchange fluxes.

Subsequently, about 5 h after inoculation, the PP pathway is activated. In this pathway important precursors for amino acid and polymer biosynthesis are produced. However, the production of these amino acids and macromolecules requires energy, which is provided by glycolysis and TCA cycle. This explains why the cells first prioritizes energy generation and starts producing macromolecules later. These amino acids contribute to the growth as well, leading to an even increased growth rate. In the exponential growth phase, the glycolysis, TCA cycle and PP pathway are all active.

This optimal temperature experiment leads to the identification of active metabolic pathways during different phases of microbial growth. This knowledge is now used to study the metabolic activity during an induced lag phase.

### *3.2.4. Results for the temperature shift experiment*

Finally, also the flux estimation for the temperature shift experiment is described. Also in this case, the null space basis was estimated from the data, together with the spline parameters and knots. This resulted in spline knot locations located at 4.41, 4.50, 5.50, 6.51, 9.00 and 10.00 h for the first free flux, 2.95, 6.00, 9.00 and 9.13 h for the second, 7.43 and 8.00 h for the third and 5.00, 7.00, 9.00 and 10.26 h for the last free flux. The estimated outputs and fluxes will now be discussed.

*Goodness of fit with respect to measurements.* The full set of measurements along with the simulated outputs of the dynamic model after the estimation are given in Figure 7. Again, also in this case, a satisfactory fit is obtained. The only real discrepancy between the data and the model is in the biomass concentration at the end of the experiment, where the model slightly underestimates the data. This should be taken into account when drawing conclusions from the estimates, but as the difference lies at the end of the experimental horizon, there will be no implications on the results regarding the induced lag phase. When the reduced chi-squared measure is computed for this experiment, a value of 6.93 is obtained. Also for this experiment, this value indicates that not all data is completely captured which is for this experiment in line with the observations of e.g., the biomass evolution.

*Analysis of metabolic pathways.* The flux estimates for the different pathways are given in Figures 9 to 15. As the experimental horizon is started in the middle of an exponential growth phase, the fluxes in the beginning of the experiment are not zero, but very low. The growth rate is low, as the temperature is far below optimal. The glycolysis (Figure 9), TCA cycle (Figure 11) and PP pathway (Figure 13) are simultaneously active, but at a very low level ($\leq$ 0.02 mmol/L.h) due to the suboptimal temperature. As soon as the temperature shift is imposed, the flux distribution changes and the glycolysis and TCA cycle gain activity along with the exchange rates. As in the optimal temperature experiment, the PP pathway only becomes active 2 h after the TCA cycle and glycolysis have been activated. Subsequently, they all go through a maximum, after which they quite sharply return to a very low or inactive level, marking the beginning of the stationary phase.

This change in activity of metabolic pathways during the induced lag phase is possibly a characteristic response. The conclusions from the experiment at optimal temperature can be extended to this experiment. As soon as the environment of the cells changes, for example by this sudden temperature shift, they behave similarly to the phase of initial growth during an experiment at constant temperature. Energy generation seems to be prioritized as a response to the induced lag phase, and the TCA cycle becomes more active. As the cells gain energy, the PP pathway is reactivated, while the TCA cycle activity decreases again. Growth proceeds similarly to the experiment at optimal temperature. Unfortunately, the data available before the induced lag phase is not optimal due to the low temperature. The cells were not in a fast exponential growth phase (with fast biomass growth and acid secretion) before the induced lag phase. A response to the temperature profile can be observed, but the dynamics before this induced lag phase are not well known.

## 4. Conclusions

The goal of this research is to validate a novel algorithm which estimates the dynamic fluxes in an induced lag phase in predictive microbiology. The optimization procedures were carried out on data obtained from different experiments, one at the optimal growth temperature and one with an imposed temperature profile. The analysis revealed some interesting patterns in metabolic activity during different phases of growth.

Initially, energy or ATP production is prioritized by activity of the glycolysis and TCA cycle pathways. This happens because energy is required for activation of the more complex processes and metabolic pathways. Oxygen is required to keep the TCA cycle active, and $CO_2$ is released as a result of the carbon balance. Excess intermediates such as acetate are secreted. Subsequently, the metabolic pathways for precursor and

amino acid production become active. These pathways require the energy provided by the TCA cycle. Therefore, they only become more active later in the growth process, when the culture is growing exponentially.

A similar behavior is noticed for the induced lag phase after a temperature shift. The cells respond to this environmental change by activation of energetic pathways and shutting down the PP pathway, which is responsible for important metabolic precursor production.The generation of ATP is prioritized and glucose is used as an energy source. This behavior makes sense, as the new environment allows an increased growth rate, for which energy is required. Once the cells re-enter the exponential growth phase, the PP pathway is activated again, combined with amino acid production. Glucose is used as both an energy source and metabolite production source.

This response may be characteristic for the behavior of cells in an induced lag phase, however, more experimental research might be required to confirm this assumption. The focus of this contribution has been to investigate the capability of the developed methodology to describe the time-varying metabolic activity of a batch culture. All steps necessary to set up such a study have been described in this contribution, and the estimates show promising results for its future applicability.

Furthermore, it is suggested in future research to reduce the uncertainty in the measurements. This is in particular important to determine the source of the unexplained behavior. In the presented results, the difference between the observed and estimated behavior can be attributed to two different sources: the large uncertainty in the measurements or a potential lack of model accuracy. At the moment it is not yet possible to discriminate between the two potential causes.

Furthermore, there are two ways the authors envision the use of the described methodology in predictive microbiology in foods. On the one hand, dMFA can be used to estimate the fluxes during relevant experiments, as it was used in this contribution, and these estimated fluxes can then be used to gain more insight on the intracellular level, which can be integrated into existing macroscopic model structures. On the other hand, the dMFA model structure can also be integrated directly into the widely used macroscopic models, resulting in a next generation predictive model that can be used in a wider range of conditions, see [11] for an application where the network model is employed to optimize growth in a bioreactor.

## Acknowledgments

The research was supported by: PFV/10/002 (OPTEC), FWO KAN2013 1.5.189.13, FWO-G.0930.13 and BelSPO: IAP VII/19 (DYSCO). DT holds PDM grant 2015/134 from KU Leuven.

548 **List of symbols and abbreviations**

| | |
|---|---|
| 3PG | 3-phosphoglycerate |
| 6PG | 6-phosphogluconate |
| AA | amino acid |
| Ac | Acetic acid |
| AcCoA | Acetyl-Coenzyme A |
| AIC | Akaike information criterion |
| $AIC_c$ | corrected AIC criterion value |
| AKG | alpha-ketoglutarate |
| ATP | adenosine triphosphate |
| Ala | Alanine |
| Arg | arginine |
| Asn | Asparagine |
| Asp | Aspartate |
| CFU | total viable plate count (colony forming units) |
| Cit | citrate |

549

| | |
|---|---|
| $c_{\text{in}}$ | vector of inlet tank concentrations |
| $c_{\text{int}}$ | vector of intracellular metabolite concentrations |
| CTR | carbon dioxide transfer rate |
| Cys | cysteine |
| $d$ | number of free fluxes |
| DHAP | dihydroxyacetone-phosphate |
| $dMFA$ | dynamic metabolic flux analysis |
| DW | dry weight |
| E4P | erythrose-4-phosphate |
| $F$ | objective function |
| $F^{in}$ | incoming gas flow rate |
| F6P | fructose-6-phosphate |
| FBP | 1,6-fructose-biphosphate |
| FTHF | formyltetrahydrofolate |
| Fum | fumarate |

| | |
|---|---|
| $g$ | number of internal spline knots |
| G3P | glyceraldehyde-3-phosphate (G3P) |
| G6P | glucose-6-phosphate |
| GAP | glyceraldehyde-3-phosphate |
| Gln | glutamine |
| Glu | glutamate |
| Gluc | glucose |
| Gly | glycine |
| His | histidine |
| HPLC | high performance liquid chromatography |
| Icit | isocitrate |
| Ile | isoleucine |
| $\mathbf{I}$ | identity matrix |
| $\mathbf{I}_{\text{irr}}$ | irreversibility matrix |
| $\mathbf{K}$ | basis for the null space of the intracellular stoichiometric matrix |
| $k$ | spline degree |
| $l$ | minimum number of time points for starting the free flux estimation |
| L-DAP | L-diaminopimelate |
| Leu | Leucine |
| Lys | Lysine |
| Mal | malate |
| METHF | Methyltetrahydrofolate |
| MEETHF | methylentetrahydrofolate |
| Met | Methionine |
| $m$ | total number of metabolites/concentration states |
| $m_{ext}$ | number of extracellular metabolites |
| $m_{ij}$ | average measurement for output $j$ at time point $t_i$ |
| $m_int$ | number of intracellular metabolites |
| $n$ | total number of reactions |
| NAD | nicotinamide adenine dinucleotide |
| NADPH | nicotinamide adenine dinucleotide phosphate |

| | |
|---|---|
| $n_g$ | total number of spline knots over all free fluxes |
| $n_{\text{irr}}$ | number of irreversible reactions |
| $n_{\text{meas}}$ | number of measurements |
| $n_{\text{out}}$ | number of outputs |
| $n_p$ | number of parameters |
| $n_{\text{rev}}$ | number of reversible reactions |
| $n_{time}$ | number of measurement time points |
| $n_x$ | number of states |
| $n_y$ | number of outputs |
| OAC | oxaloacetate |
| OUR | oxygen uptake rate |
| $\mathbf{p}$ | parameter vector |
| Phe | phenylalanine |
| Pro | proline |
| P5P | pentose-5-phosphate |
| PEP | phosphoenolpyruvate |
| PP | pentose phosphate |
| Pyr | pyruvate |
| $p_u$ | vector of spline parameters |
| $\mathbf{q}_{\text{bio}}$ | biomass selection vector |
| R5P | ribose-5-phosphate |
| RID | refractive index detector |
| Ru5P | Ribulose-5-phosphate |
| $\mathbf{S}$ | full stoichiometric matrix |
| S7P | seduheptulose-7-phosphate |
| Ser | Serine |
| Suc | succinate |
| SucCoA | succinyl-Coenzyme A |
| $\mathbf{s}_{\text{bio}}$ | row of the stoichiometric matrix corresponding to the biomass pseudometabolite |
| $\mathbf{S_e}$ | combined extracellular and biomass stoichiometric matrix |
| $\mathbf{S_{ext}}$ | rows of the stoichiometric matrix corresponding to extracellular metabolites |

| | |
|---|---|
| $\sigma_{ij}$ | measurement standard deviation for output $j$ at time point $t_i$ |
| $\mathbf{S_{int}}$ | rows of the stoichiometric matrix corresponding to intracellular metabolites |
| $T$ | temperature |
| TA | transaldolase |
| TCA | tricarboxylic acid |
| Thr | Threonine |
| TK | transketolase |
| Trp | tryptophane |
| Tyr | tyrosine |
| $t_0$ | initial time |
| $t_f$ | final time |
| $t_{knot}$ | vector containing the knot locations for each free flux |
| $\mathbf{u}$ | vector of free fluxes |
| $V$ | reactor volume |
| $V_m$ | ideal gas standard molar volume |
| $\mathbf{x}$ | vector of (concentration) states |
| $\mathbf{x_0}$ | vector of initial values for the states |
| $\mathbf{x_{ext}}$ | vector of extracellular metabolite concentrations |
| $\mathbf{K}_{\text{gas}}$ | gases null space basis matrix |
| $x_{\text{head}}$ | vector of headspace molar fractions |
| $x_{\text{head,in}}$ | vector of inlet gas mole fractions |
| $x^{in}_{\text{CO}_2}$ | incoming air molar fraction of carbon dioxide |
| $x^{in}_{\text{O}_2}$ | incoming air molar fraction of oxygen |
| $x^{out}_{\text{CO}_2}$ | outgoing air molar fraction of carbon dioxide |
| $x^{out}_{\text{O}_2}$ | outgoing air molar fraction of oxygen |
| $\mathbf{x}_{\text{sel}}$ | vector of states selected for measurement |
| Val | Valerate |
| X5P | xylulose-5-phosphate |
| $\mathbf{y}$ | vector of model outputs |
| $\mathbf{y_m}$ | vector of measured values for the outputs |
| $\mathbf{z}$ | vector of algebraic states |

## References

[1] Andersson, J., Åkesson, J., Diehl, M., 2012. CasADi – A symbolic package for automatic differentiation and optimal control. In: Forth, S., Hovland, P., Phipps, E., Utke, J., Walther, A. (Eds.), Recent Advances in Algorithmic Differentiation. Vol. 87 of Lecture Notes in Computational Science and Engineering. Springer, Berlin Heidelberg, pp. 297–307.

[2] Antoniewicz, M. R., 2013. Dynamic metabolic flux analysis - tools for probing transient states of metabolic networks. Current Opinion in Biotechnology 24, 973–978.

[3] Antoniewicz, M. R., Kraynie, D. F., Laffend, L. A., Gonzalez-Lergier, J., Kelleher, J. K., Stephanopoulos, G., 2007. Metabolic flux analysis in a nonstationary system: Fed batch fermentation of a high yielding strain of *E. coli* producing 1,3-propanediol. Metabolic Engineering 9, 277–292.

[4] Baert, K., Valero, A., De Meulenaer, B., Samapundo, S., Ahmed, M. M., Bo, L., Debevere, J., Devlieghere, F., 2007. Modeling the effect of temperature on the growth rate and lag phase of penicillium expansum in apples. International Journal of Food Microbiology 118(2), 139–150.

[5] Beste, D., Noh, K., Niedenfuhr, S., Mendum, T., Hawkins, N. , Ward, J., Beale, M., Wiechert, W., McFadden., J., 2013. 13C-flux spectral analysis of host-pathogen metabolism reveals a mixed diet for intracellular mycobacterium tuberculosis. Chemistry & Biology 20, 1012–1021.

[6] Berthoumieux, S., Brilli, M., de Jong, H., Kahn, D., Cinquemani, E., 2011. Identification of metabolic network models from incomplete high-throughput datasets. Bioinformatics 27(13), 186–195.

[7] Bevington, P.R and Robinson, D.K., 2003, Data reduction and error analysis for the physical sciences; 3rd ed., McGraw-Hill, New York, USA.

[8] Biegler, L.T., 2010, Nonlinear Programming: Concepts, Algorithms, and Applications to Chemical Processes, MPS-SIAM, Philadelphia, USA.

[9] Brul, S., Westerhoff, H., 2007. Systems biology and food science. In: Brul, S., van Gerwen, S., Zwietering, M. (Eds.), Modelling microorganisms in food. Woodhead, pp. 250–288.

[10] Burnham, K.P. and Anderson, D.R., 2004. Multimodel Inference: Understanding AIC and BIC in Model Selection, Modelling microorganisms in food. Sociological Methods & Research 33(2), 261–304.

[11] Chang, L., Liua, X., Henson, M., 2016. Nonlinear model predictive control of fed-batch fermentations using dynamic flux balance models. Journal of Process Control 42, 137–149.

[12] Delignette-Muller, M. L., Baty, F., Cornu, M., Bergis, H., 2005. Modelling the effect of a temperature shift on the lag phase duration of *Listeria monocytogenes*. International Journal of Food Microbiology 100(1-3), 77–84.

[13] de Boor, C., 2001. A Practical Guide to Splines. Springer-Verlag, New York, USA.

[14] Dierckx, P., 1993. Curve and surface fitting with splines. Oxford Univ. Press, New York, USA.

[15] Emmerling, M., Dauner, M., Ponti, A., Fiaux, J., Hochuli, M., Szyperski, T., Wüthrich, K., Bailey, J., Sauer, U., 2002. Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*. Journal of Bacteriology 184, 152–164.

[16] Francois, K., Devlieghere, F., Standaert, A. R., Geeraerd, A. H., Van Impe, J. F., Debevere, J., 2006. Effect of environmental parameters (temperature, pH and aw) on the individual cell lag phase and generation time of *Listeria monocytogenes*. International Journal of Food Microbiology 108(3), 326–335.

[17] Kayser, A., Weber, J., Hecht, V., Rinas, U., 2005. Metabolic flux analysis of *Escherichia coli* in glucose-limited continuous culture. I. Growth rate dependent metabolic efficiency at steady state. Microbiology 151, 693–706.

[18] Kuhl, P., Diehl, M., Kraus, T., Schlöder, J., Bock, H.G., 2011. A real-time algorithm for moving horizon state and parameter estimation. Computers & Chemical Engineering 35, 71–83.

[19] Mahadevan, R., Edwards, J., Doyle, F., 2002. Dynamic flux balance analysis of diauxic growth in *Escherichia coli*. Biophysical Journal 83 (3), 1331–1340.

[20] Martinez, V., Buchsteiner, M., Gray, P., Nielsen, L., and Quek, L.-E., 2015. Dynamic metabolic flux analysis using B-splines to study the effects of temperature shift on CHO cell metabolism. Metabolic Engineering Communications 2, 46–57.

[21] McMeekin, T. A., Bowman, J., McQuestin, O., Mellefont, L., Ross, T., Tamplin, M., 2008. The future of predictive microbiology: Strategic research, innovative applications and great expectations. International Journal of Food Microbiology 128, 2–9.

[22] McMeekin, T. A., Brown, J., Krist, K., Miles, D., Neymeyer, K., Nichols, D. S., Olley, J., Presser, K., Ratkowsky, D. A., Ross, T., Salter, M., Soontranon, S., 1997. Quantitative microbiology: A basis for food safety. Emerging Infectious Diseases 3(4), 541–549.

608 [23] Mellefont, L. A., Ross, T., 2003. The effect of abrupt shifts in temperature on the lag phase duration of
609 *Escherichia coli* and *Klebsiella oxytoca*. International Journal of Food Microbiology 83, 295–305.

610 [24] Metris, A., George, S., Baranyi, J., 2012. Modelling osmotic stress by flux balance analysis at the genomic
611 scale. International Journal of Food Microbiology 152(3), 123–128.

612 [25] Muñoz-Cuevas, M., Fernández, P. S., George, S., Pin, C., 2010. Modeling the lag period and exponential
613 growth of *Listeria monocytogenes* under conditions of fluctuating temperature and water activity values.
614 Applied and Environmental Microbiology 76(9), 2908–2915.

615 [26] Nanchen, A., Schicker, A., Sauer, U., 2006. Nonlinear dependenc of intracellullar fluxes on growth rate
616 in miniaturized continuous cultures of *Escherichia coli*. Applied and Environmental Microbiology 72,
617 1164–1172.

618 [27] Swinnen, I., Bernaerts, K., Dens, E., Geeraerd, A., Van Impe, J., 2004. Predictive modelling of the
619 microbial lag phase: a review. International Journal of Food Microbiology 94, 137–159.

620 [28] Swinnen, I. A. M., Bernaerts, K., Gysemans, K., Van Impe, J. F., 2005. Quantifying microbial lag
621 phenomena due to a sudden rise in temperature: a systematic macroscopic study. International Journal
622 of Food Microbiology 100, 85–96.

623 [29] Van Impe, J. F., Vercammen, D., Van Derlinden, E., 2013. Toward a next generation of predictive
624 models: A systems biology primer. Food Control 29(2), 336–42.

625 [30] Vercammen, D., Logist, F., Van Impe, J., 2014. Dynamic estimation of specific fluxes in metabolic
626 networks using non-linear dynamic optimization. BMC Systems Biology 8, 132.

627 [31] Vercammen, D., Logist, F., Van Impe, J., 2014. Estimation of specific fluxes in metabolic networks using
628 non-linear dynamic optimization. In: Klemes, J., Friedler, F., Mizsey, P. (Eds.), Proceedings of the 24th
629 European Symposium of Computer Aided Process Engineering. Budapest, Hungary, pp. 289–294.

630 [32] Vercammen, D., Logist, F., Van Impe, J., 2016. Online moving horizon estimation of fluxes in metabolic
631 reaction networks. Journal of Process Control 37, 1–20.

632 [33] von Stockar, U., Liu, J., 1999. Does microbial life always feed on negative entropy? Thermodynamic
633 analysis of microbial growth. Biochimica et Biophysica Acta - Bioenergetics 1412, 191–211.

634 [34] Wächter, A., Biegler, L. T., 2006. On the implementation of a primal-dual interior point filter line search
635 algorithm for large-scale nonlinear programming. Mathematical Programming 106(1), 25–57.

636 [35] Whiting, R. C., Bagi, L. K., 2002. Modeling the lag phase of *Listeria monocytogenes*. International
637 Journal of Food Microbiology 73, 291–295.

638 [36] Wiechert, W., 2001. 13C metabolic flux analysis. Metabolic Engineering 3(3), 195–206.

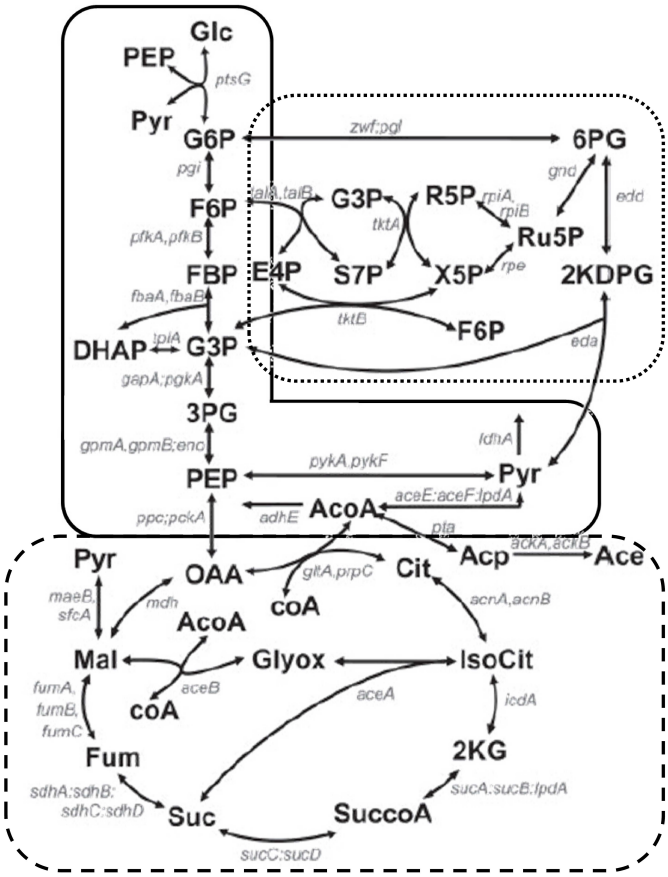


Figure 1: **Simplified metabolic reaction network.**Reaction network of *E. coli* central carbon metabolism [6]. The glycolysis pathway is indicated by the solid box, the pentose phosphate pathway by the box in dotted lines and the tricarboxylic acid cycle by the box in dashed lines.

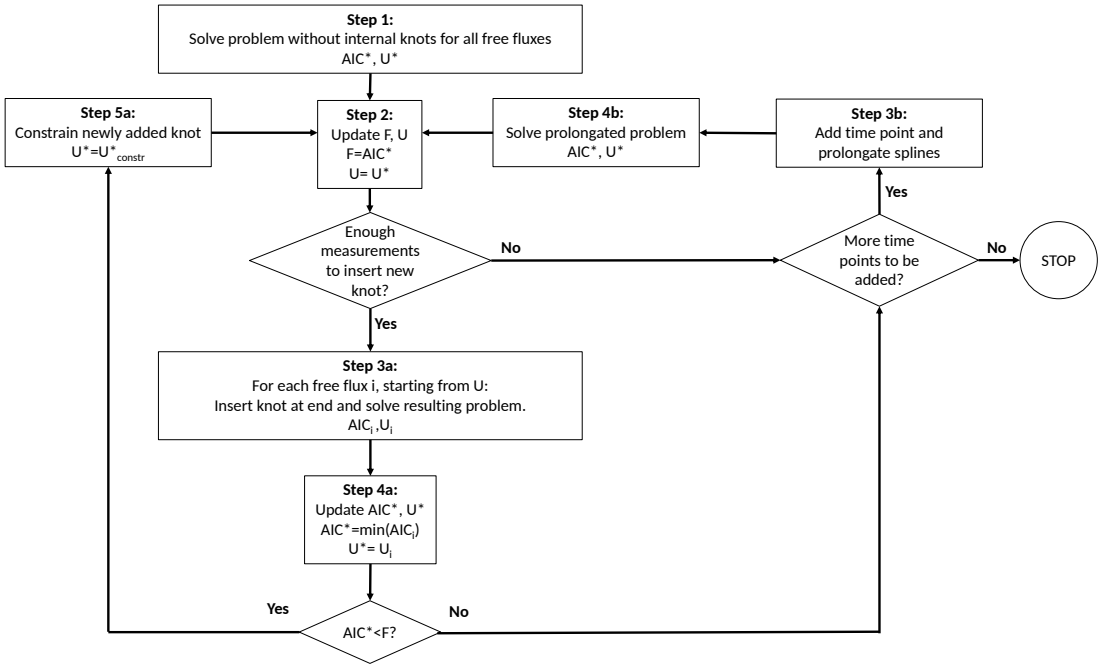


Figure 2: **Schematic algorithm description.** Schematic description of the employed incremental free flux estimation algorithm (adapted from [30]).

Table 1: **States in the dMFA model.** Overview of the states in the dMFA model, with inlet concentrations for the headspace states. The checkmarks indicate whether a state is measured or not.

| Metabolite | Measured | Inlet conc. |
|---|---|---|
| **Metabolite states x** | | |
| AKG | ✓ | |
| Ace | ✓ | |
| Amm | ✓ | |
| Sulph | | |
| Glu | ✓ | |
| $O_2$ | | |
| $CO_2$ | | |
| Bio | ✓ | |
| **Headspace states $\mathbf{x}_{\text{head}}$** | | |
| $O_2$ | ✓ | 0.2097 |
| $CO_2$ | ✓ | 0.0004 |



Journal homepage: https://www.journals.elsevier.com/food-and-bioproducts-processing/
Original file available at: http://www.sciencedirect.com/science/article/pii/S0960308516301110

Postprint version of paper published in Food and Bioproducts Processing 2017, vol. 102, p. 1-19.
**The content is identical to the published paper, but without the final typesetting by the publisher.**

Table 2: **Parameter values.** Model parameter values for the dMFA model.

| Parameter | Symbol | Value | Units |
|---|---|---|---|
| Bioreactor medium volume | $V_{\text{liq}}$ | 3.5 | L |
| Reciprocal of ideal gas molar volume | $M$ | 40.82 | mmol/L |
| Inlet air flow rate | F | 120.0 | L/h |

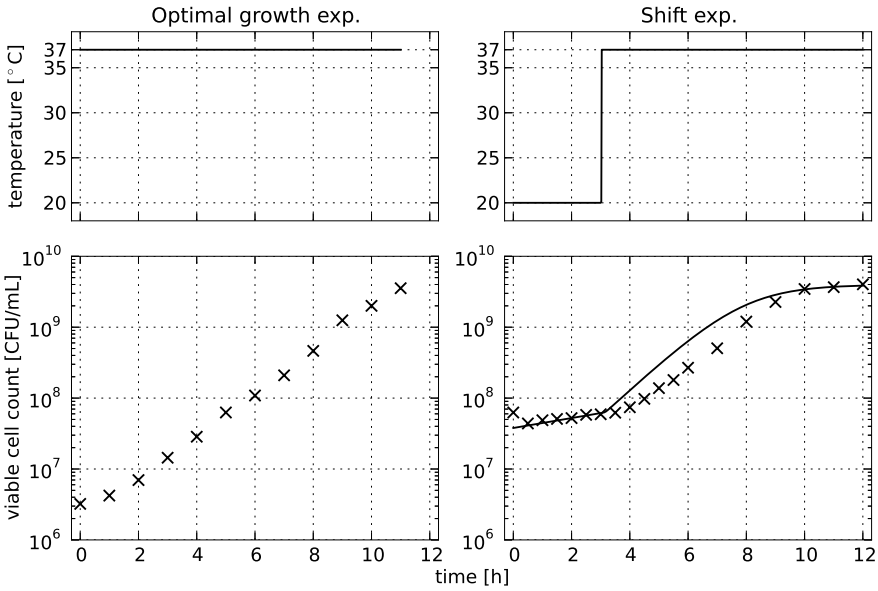


Figure 3: **Temperature profiles and growth curves.** Temperature profiles (top) and biomass growth curves (bottom) for the experiment with optimal growth temperature (left) and with a temperature shift (right). A simulation with the same exponential growth rates, but without intermediate lag phase is shown by the solid line in the bottom right figure, illustrating the presence of an induced lag phase in the experimental data.



Journal homepage: https://www.journals.elsevier.com/food-and-bioproducts-processing/
Original file available at: http://www.sciencedirect.com/science/article/pii/S0960308516301110

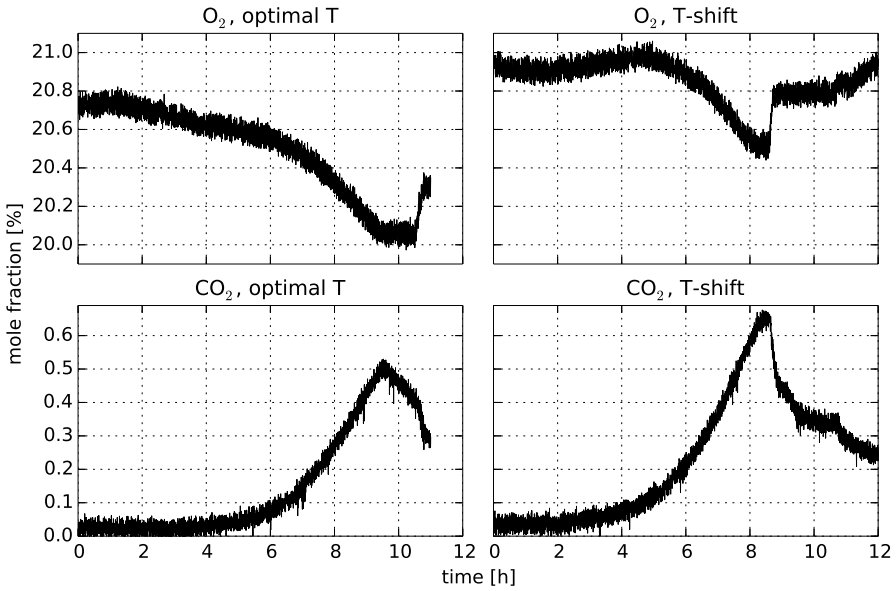


Figure 4: **Headspace gas concentrations.** The concentration of $O_2$ (on the top) and $CO_2$ (on the bottom) in the headspace of the bioreactor during the experiment at optimal growth temperature (on the left) and the temperature shift experiment (on the right).

Table 3: Glycolysis fluxes.

| Flux | Reaction |
|---|---|
| flux 1 | 1 G6P $\leftrightarrow$ 1 F6P |
| flux 2 | 1 F6P + 1 ATP $\rightarrow$ 1 FBP |
| flux 3 | 1 FBP $\leftrightarrow$ 1 DHAP + 1 GAP |
| flux 4 | 1 DHAP $\leftrightarrow$ 1 GAP |
| flux 5 | 1 GAP $\leftrightarrow$ 1 3PG + 1 ATP + 1 NADH |
| flux 6 | 1 3PG $\leftrightarrow$ 1 PEP |
| flux 7 | 1 PEP $\rightarrow$ 1 Pyr + 1 ATP |
| flux 17 | 1 Pyr $\rightarrow$ 1 AcCoA + 1 $CO_2$ + 1NADH |

Postprint version of paper published in Food and Bioproducts Processing 2017, vol. 102, p. 1-19.
**The content is identical to the published paper, but without the final typesetting by the publisher.**

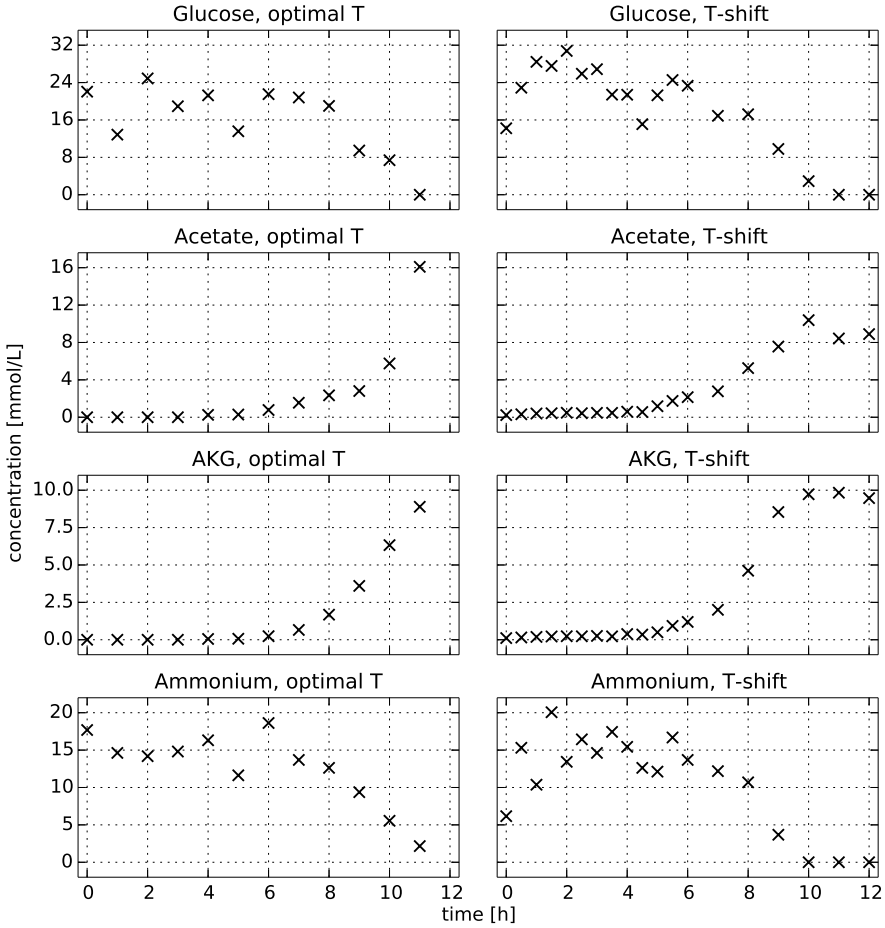


Figure 5: **Extracellular metabolite concentrations.** Concentrations of extracellular metabolites for the experiment at the optimal growth temperature (on the left) and during the temperature shift experiment (on the right), measured using HPLC or testkits.

Table 4: Tricarboxylic acid cycle fluxes.

| Flux | Reaction |
|---|---|
| flux 18 | 1 OAC + 1 AcCoA → 1 Cit |
| flux 19 | 1 Cit ↔ 1 ICit |
| flux 20 | 1 ICit ↔ 1 AKG + 1 $CO_2$ + 1 NADPH |
| flux 21 | 1 AKG → 1 SucCoA + 1 $CO_2$ + 1 NADH |
| flux 22 | 1 SucCoA ↔ 1 Suc + 1 ATP |
| flux 23 | 1 Suc ↔ 1 Fum + 1 $FADH_2$ |
| flux 24 | 1 Fum ↔ 1 Mal |
| flux 25 | 1 Mal ↔ 1 OAC + 1 NADH |
| flux 26 | 1 PEP + 1 $CO_2$ → 1 OAC |
| flux 27 | 1 AcCoA ↔ 1 Ac + 1 ATP |
| flux 52 | 1 NADH + 0.5 $O_2$ → 3 ATP |
| flux 53 | 1 $FADH_2$ + 0.5 $O_2$ → 2 ATP |



Journal homepage: https://www.journals.elsevier.com/food-and-bioproducts-processing/
Original file available at: http://www.sciencedirect.com/science/article/pii/S0960308516301110

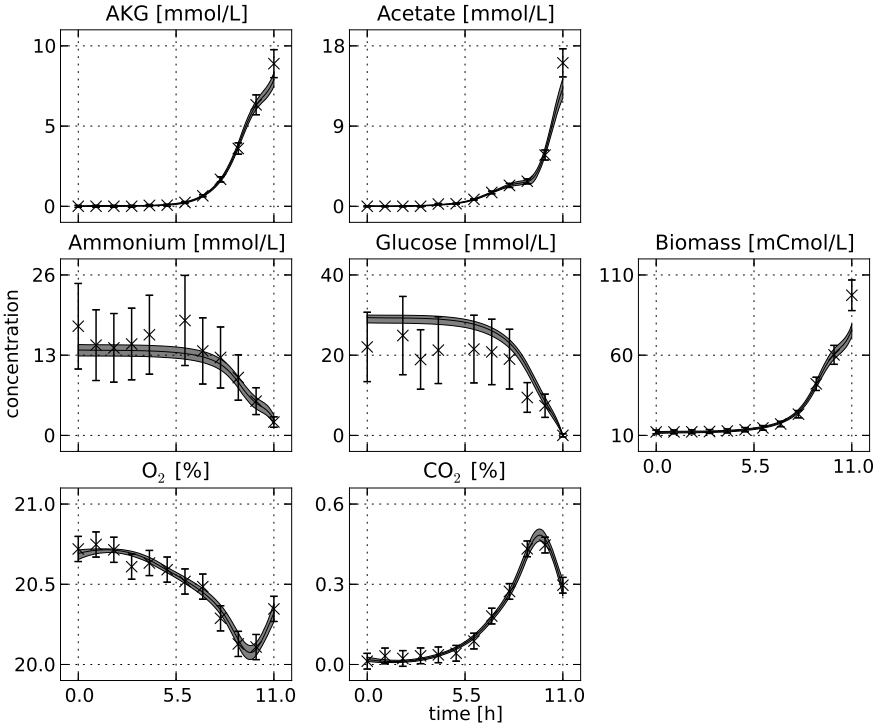


Figure 6: **Dataset and fitted state evolutions for experiment at optimal growth temperature.** Fitted measurements for the optimal temperature experiment. The 95% confidence intervals for the estimated outputs are shown in gray, also the measurement error bars denote a 95% uncertainty interval.

Table 5: Pentose phosphate pathway fluxes.

| Flux | Reaction |
|---|---|
| flux 8 | 1 G6P $\rightarrow$ 1 6PG + 1 NADPH |
| flux 9 | 1 6PG $\rightarrow$ 1 Ru5P + 1 $CO_2$ + 1 NADPH |
| flux 10 | 1 Ru5P $\leftrightarrow$ 1 X5P |
| flux 11 | 1 Ru5P $\leftrightarrow$ 1 R5P |
| flux 12 | 1 X5P $\leftrightarrow$ 1 TK-C2 + 1 GAP |
| flux 13 | 1 F6P $\leftrightarrow$ 1 TK-C2 + 1 E4P |
| flux 14 | 1 S7P $\leftrightarrow$ 1 TK-C2 + 1 R5P |
| flux 15 | 1 F6P $\leftrightarrow$ 1 TA-C3 + 1 GAP |
| flux 16 | 1 S7P $\leftrightarrow$ 1 TA-C3 + 1 E4P |

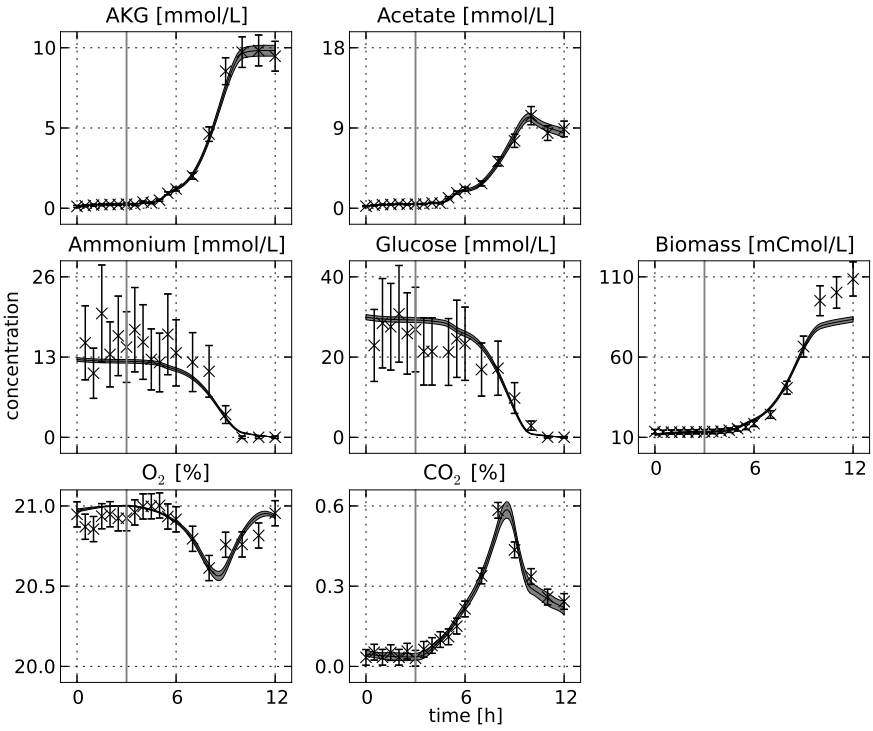


Figure 7: **Dataset and fitted state evolutions for experiment for temperature shift experiment.** Fitted measurements for the temperature shift experiment. The 95% confidence intervals for the estimated outputs are shown in gray, the measurement error bars denote a 95% uncertainty interval. The time where the temperature shift was performed is indicated with a gray vertical line, i.e., at 3 h.

Table 6: Exchange fluxes.

| Flux | Reaction |
|---|---|
| flux 55 | 1 Gluc[e] + 1 ATP → 1 G6P |
| flux 56 | 1 AKG → 1 AKG[e] |
| flux 57 | 1 Ac ↔ 1 Ac[e] |
| flux 58 | 1 $CO_2$ → 1 $CO_2$[e] |
| flux 59 | 1 $O_2$[e] → 1 $O_2$ |
| flux 60 | 1 $NH_3$[e] → 1 $NH_3$ |
| flux 61 | 1 $SO_4$[e] → 1 $SO_4$ |
| flux 62 | 0.488 Ala + 0.281 Arg + 0.229 Asn + 0.229 Asp + 0.087 Cys + 0.250 Glu + 0.250 Gln + 0.582 Gly + 0.090 His + 0.276 Ile + 0.428 Leu + 0.326 Lys + 0.146 Met + 0.176 Phe + 0.210 Pro + 0.205 Ser + 0.241 Thr + 0.054 Trp + 0.131 Tyr + 0.402 Val + 0.205 G6P + 0.071 F6P + 0.754 R5P + 0.129 GAP + 0.619 3PG + 0.051 PEP + 0.083 Pyr + 2.510 AcCoA + 0.087 AKG + 0.340 OAC + 0.443 MEETHF + 33.247 ATP + 5.363 NADPH → 39.68 BIOMASS + 1.455 NADH |

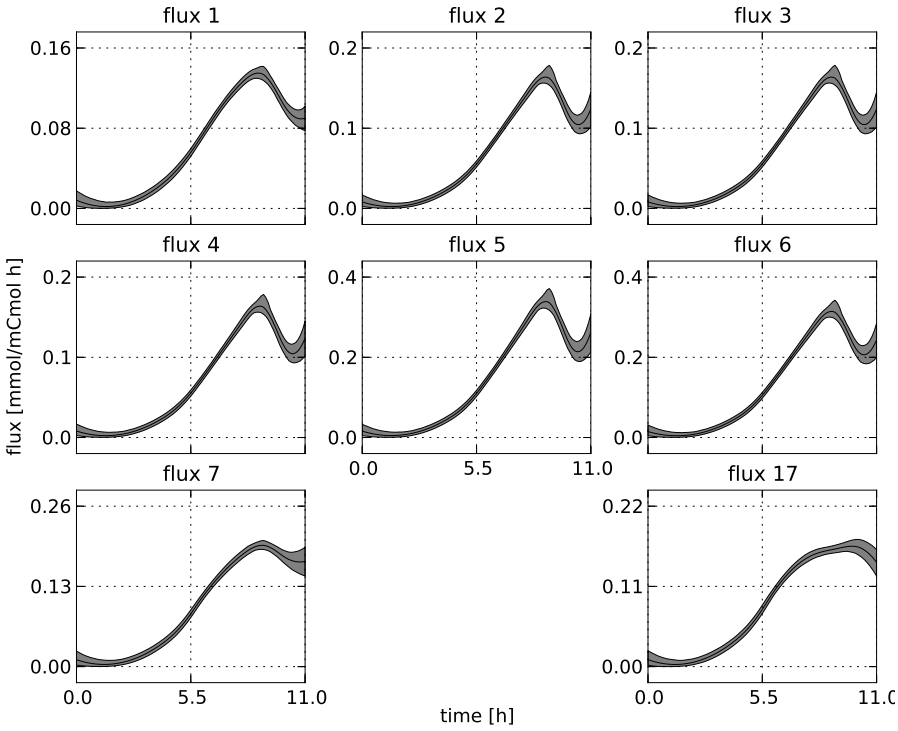


Figure 8: **Estimated glycolysis fluxes.** Estimated flux profiles for the optimal temperature experiment, for the glycolysis fluxes. The flux numbers correspond to the reaction numbers in the Supplementary data. The 95% confidence intervals for the estimated outputs are shown in gray. These fluxes are related with the ones depicted in Table 3.

Postprint version of paper published in Food and Bioproducts Processing 2017, vol. 102, p. 1-19.
**The content is identical to the published paper, but without the final typesetting by the publisher.**

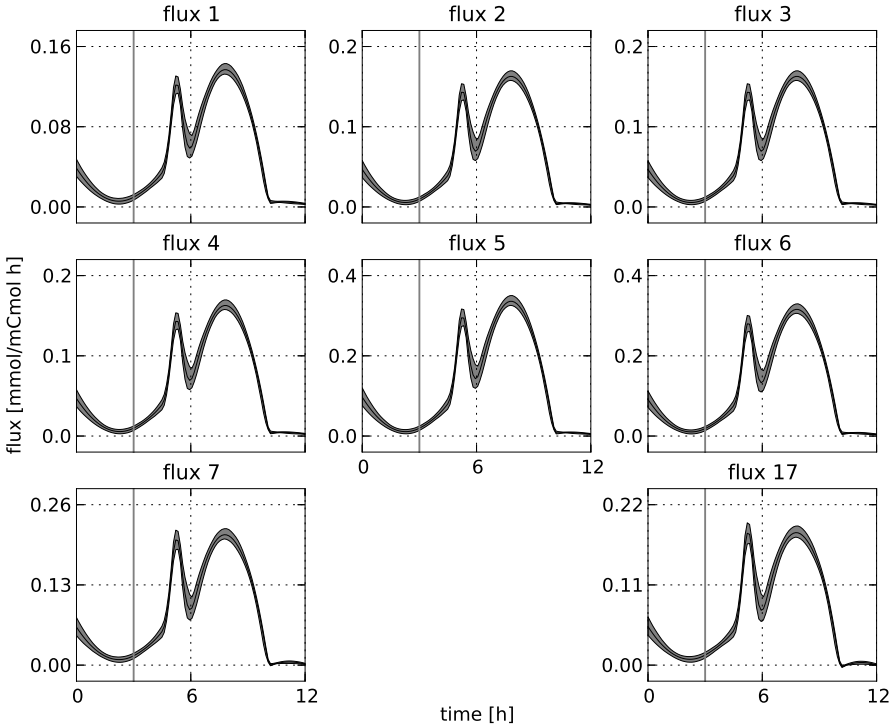


Figure 9: **Estimated glycolysis fluxes.** Estimated flux profiles for the temperature shift experiment, for the glycolysis fluxes. The flux numbers correspond to the reaction numbers in the Supplementary data. The time where the temperature shift was performed is indicated with a gray vertical line, i.e., at 3 h. The 95% confidence intervals for the estimated outputs are shown in gray. These fluxes are related with the ones depicted in Table 3.



Journal homepage: https://www.journals.elsevier.com/food-and-bioproducts-processing/
Original file available at: http://www.sciencedirect.com/science/article/pii/S0960308516301110

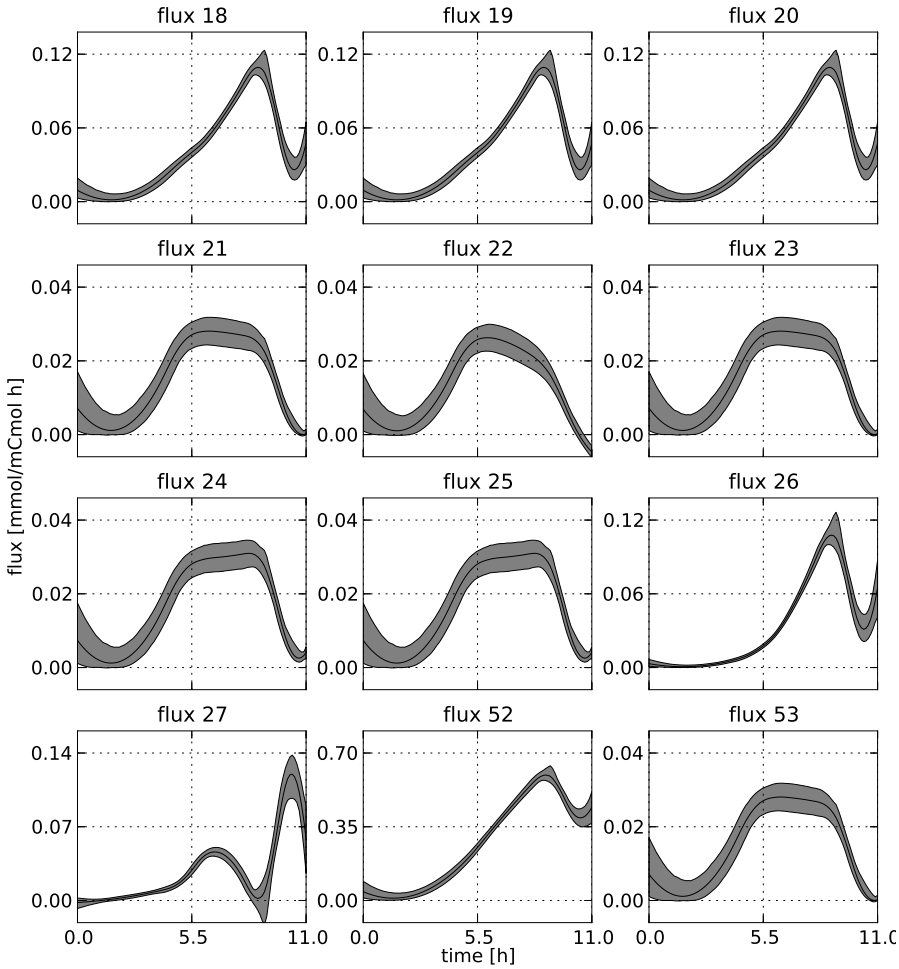


Figure 10: **Estimated TCA cycle fluxes.** Estimated flux profiles for the optimal temperature experiment, for the TCA cycle fluxes. The flux numbers correspond to the reaction numbers in the Supplementary data. The 95% confidence intervals for the estimated outputs are shown in gray. These fluxes are related with the ones depicted in Table 4.

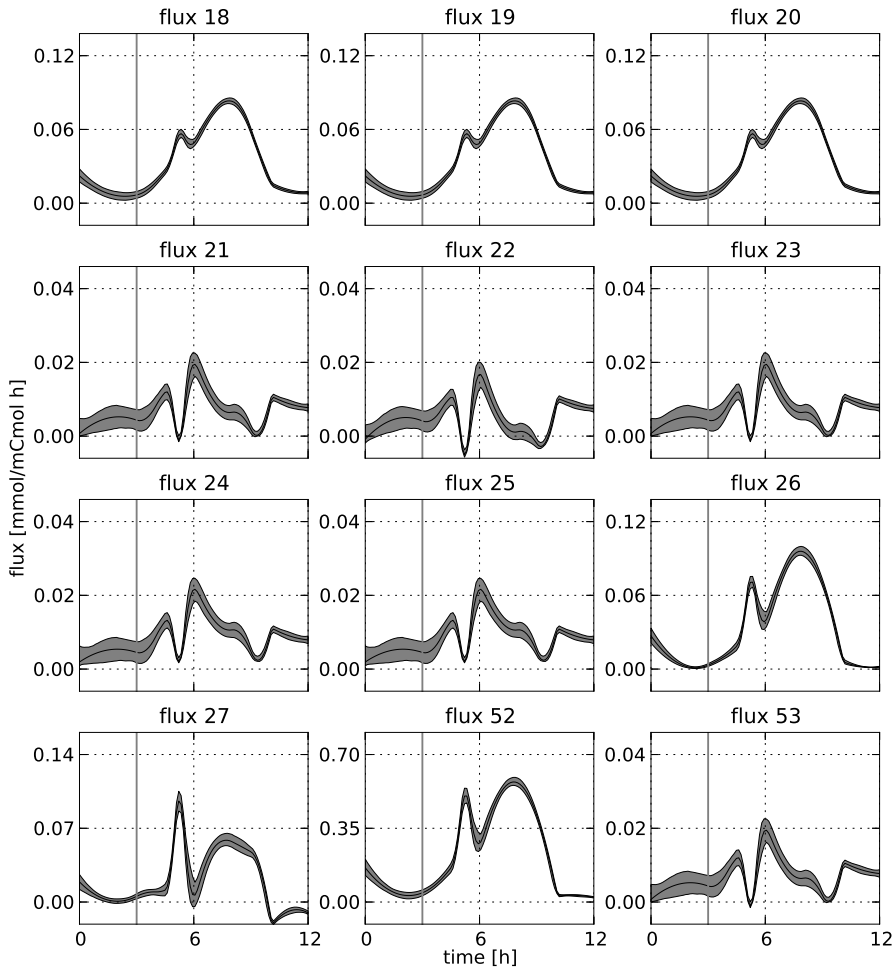


Figure 11: **Estimated TCA cycle fluxes.** Estimated flux profiles for the temperature shift experiment, for the TCA cycle fluxes. The flux numbers correspond to the reaction numbers in the Supplementary data. The time where the temperature shift was performed is indicated with a gray vertical line, i.e., at 3 h. The 95% confidence intervals for the estimated outputs are shown in gray. These fluxes are related with the ones depicted in Table 4.

Postprint version of paper published in Food and Bioproducts Processing 2017, vol. 102, p. 1-19.
**The content is identical to the published paper, but without the final typesetting by the publisher.**

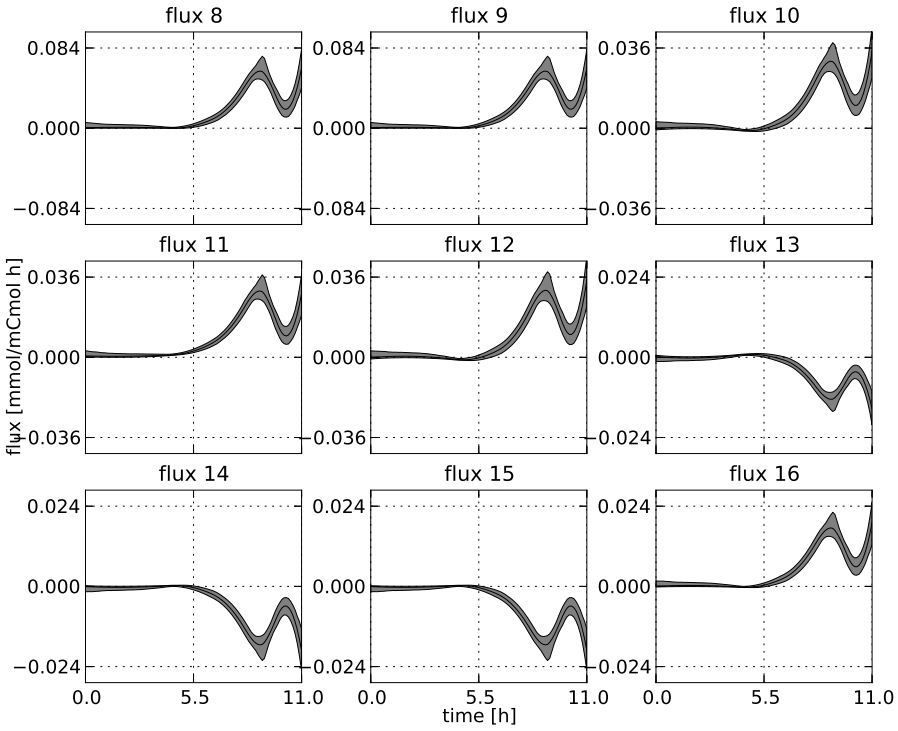


Figure 12: **Estimated PP pathway fluxes.** Estimated flux profiles for the optimal temperature experiment, for the PP pathway fluxes. The flux numbers correspond to the reaction numbers in the Supplementary data. The 95% confidence intervals for the estimated outputs are shown in gray. These fluxes are related with the ones depicted in Table 5.

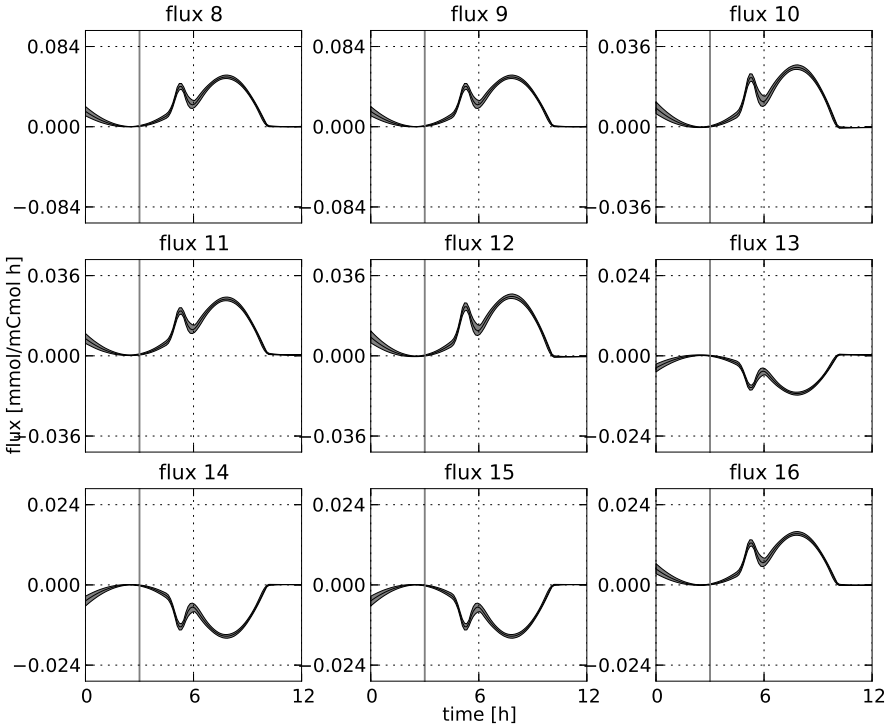


Figure 13: **Estimated PP pathway fluxes.** Estimated flux profiles for the temperature shift experiment, for the PP fluxes. The flux numbers correspond to the reaction numbers in the Supplementary data. The time where the temperature shift was performed is indicated with a gray vertical line, i.e., at 3 h. The 95% confidence intervals for the estimated outputs are shown in gray. These fluxes are related with the ones depicted in Table 5.

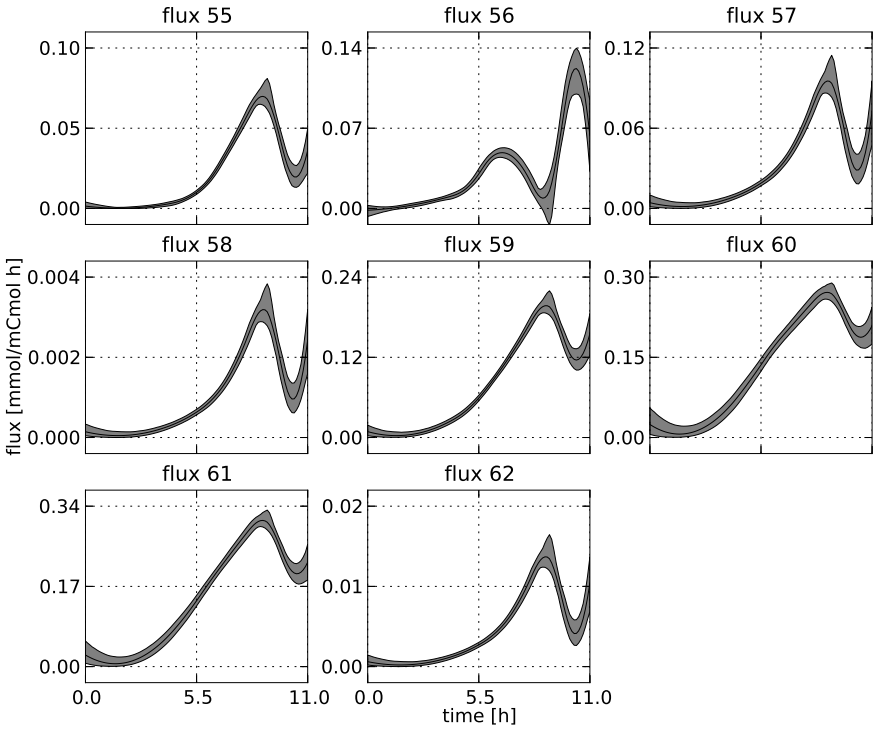


Figure 14: **Estimated exchange fluxes.** Estimated flux profiles for the optimal temperature experiment, for the exchange fluxes. The flux numbers correspond to the reaction numbers in the Supplementary data. The 95% confidence intervals for the estimated outputs are shown in gray. These fluxes are related with the ones depicted in Table 6.

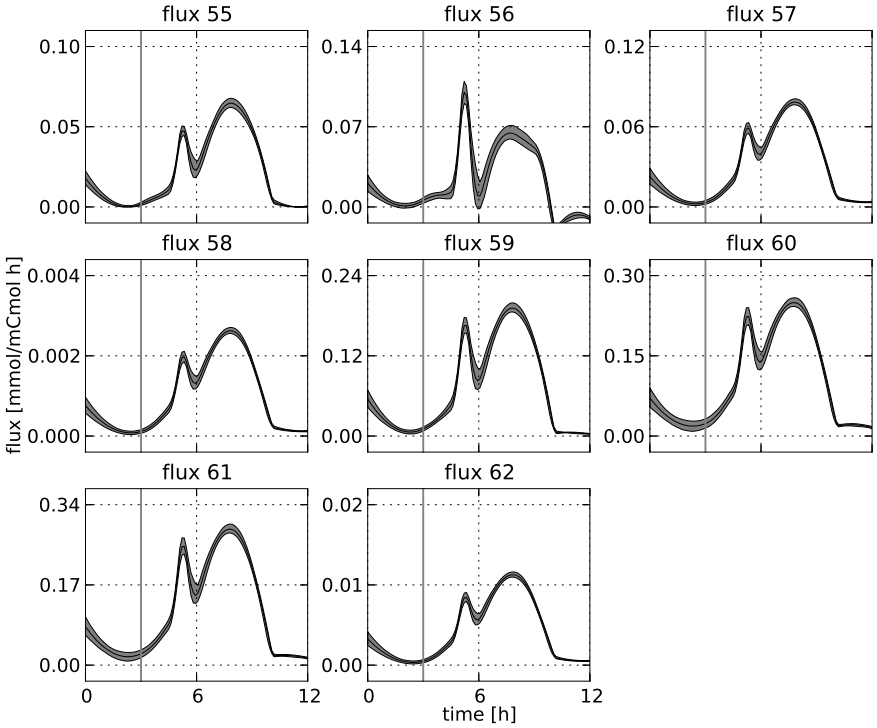


Figure 15: **Estimated exchange fluxes.** Estimated flux profiles for the temperature shift experiment, for the exchange fluxes. The flux numbers correspond to the reaction numbers in the Supplementary data. The time where the temperature shift was performed is indicated with a gray vertical line, i.e., at 3 h. The 95% confidence intervals for the estimated outputs are shown in gray. These fluxes are related with the ones depicted in Table 6.